MWT2/Illinois Campus Cluster Pilot

- Purpose
  - Show possible solution to integrate the ICC into MWT2
  - Learn how to setup Condor Flocking between two sites
  - Help determine needs for ICC Instance 2

- Illinois Campus Cluster Instance 1 (Taub)
  - HP blades with Westmere, 24GB, 160GB disk
  - InfiniBand interface to DDN for GPFS file systems
  - Tier2 will be part of ICC Instance 2 as Anchor Tenant

- 4 Taub nodes used in Pilot
  - College of Engineering loaners
  - HEP loaned 3 1TB SATA-300 drives
  - Condor master (taub208)
  - 3 Worker Nodes (taub205, taub206, taub207)

- Systems PXE boot to RAM disk
  - 24GB memory, 1GB for system, 1GB for GPFS
  - One local disk for WN Scratch, CVMFS, /tmp, etc
  - /usr/local, /home, etc located on GPFS

- Networking
  - 40Gb InifiniBand to DDN
  - No NAT available
  - Pilot Nodes connected to public network via 1Gb NIC

- IllinoisHEP Tier3gs used as Tier2
  - Gatekeeper (osgx0)
  - Condor Master (osgcondor) version 7.6.1
  - IllinoisHEP SE used for endpoints (osgx1/2/3)
  - IlliniosHEP squid (osgsquid)

- Replicate Job Environment on ICC
  - Rsync server on osgx0
  - Serve $OSGAPP, WN Client and CA
  - Must be in same directory path as IllinoisHEP
  - CVMFS installed on pilot worker nodes
  - Standard set of RPMS for Atlas releases and pilot

- Soft links tie some things together
  - /etc/grid-security/certificates → /home/osg/CA
  - /home/osgstore/tmp → /scratch.local/osgtmp

- Local users for usatlas
  - Create users usatlas1/2/3/4
  - Create groups usatlas, usatlas1/2/3/4
  - Same UID/GID as on IllinoisHEP
  - Same HOME as on IllinoisHEP
  - Avoids permission issues to OSGAPP, etc
  - Atlas pilot remember HOME when flock (.curlrc)

- Condor Flocking
  - Campus firewall block most ports to Taub
  - Condor SHARED_PORT and CCB used
  - SHARED_PORT Daemon runs on every pilot node
  - FLOCK_FROM, FLOCK_TO on each Condor
  - ALLOW_READ, ALLOW_WRITE
  - UID_DOMAIN = USAtlas
  - TRUST_UID_DOMAIN = True

- Panda Queues
  - Created two test queues for needed schedconfig differences
  - IllinoisCC and ANALY_IllinoisCC
  - Clones of IllinoisHEP and ANALY_IllinoisHEP

- Schedconfig changes

copytoolin = “None”

copytool = “lcg-cp2”

globusrsl = (condorsubmit=('+IsCCJob' TRUE)('+IsSPJob' TRUE) ('+IsProdJob' TRUE)('priority' 0)('Transfer_Executable' TRUE) ('When_To_Transfer_Output' ON_EXIT)('Should_Transfer_Files' IF_NEEDED)('Requirements' 'FileSystemDomain != \\\"\\\"'))

- IsCCJob, IsSPJob, IsProdJob (and IsAnalyJob)
  - Used by IllinoisHEP Condor to steer jobs to ICC
  - Keep from matching a job slot on IllinoisHEP
- Transfers of binaries and output between Condors
  - Atlas Pilot binary must move to ICC
  - Outputs of pilot must move back to IllinoisHEP
- FileSystemDomain != “”
  - Job will only run on node with file system of submit host
  - This allows job to run on any file system

- Job Flow with Flocking
  - Only minor difference between local and remote
  - Panda submit Atlas Pilot via Glidein to gatekeeper
  - Gatekeeper does a “condor_submit”
  - If from the CC queues, the “globusrsl” is used in submit
  - Job will not match a slot on IllinoisHEP but does on ICC
  - Atlas pilot is moved to ICC Condor and submited
  - Since UID_DOMAIN is identical, job runs as “usatlas1”
  - Atlas pilot retrieves what it needs to WN Scratch
  - Datasets fetched with lcg-cp from SE
  - Job is executed
  - Output returned to SE via lcg-cp

- Phase 1 Complete
  - Testing of production and analy from 10/5 to 10/12
  - Production went “online” 10/12
  - Analy went “online” 10/14
  - No problems seen (but only 3 worker nodes!)

- Phase 2 Next step
  - Shift from IllinoisHEP to MWT2
  - Rsync from MWT2, account UID/GID, Condor configs
  - Potential problems, MWT2 firewall, local site mover